

Review

A Tale of Native American Whole-Genome Sequencing and Other Technologies

Israel Aguilar-Ordoñez ^{1,2,*} , Josué Guzmán-Linares ³, Judith Ballesteros-Villascán ⁴ , Fernanda Mirón-Toruño ³, Alejandra Pérez-González ³ , José García-López ³, Fabricio Cruz-López ³  and Enrique Morett ^{1,*}

¹ Instituto de Biotecnología, Universidad Nacional Autónoma de México (UNAM), Cuernavaca 62210, Mexico

² Instituto Nacional de Medicina Genómica (INMEGEN), Mexico City 14610, Mexico

³ School of Biotechnology Puebla de Zaragoza, Benemérita Universidad Autónoma de Puebla (BUAP), Puebla 72000, Mexico

⁴ Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO), Centro de Investigación Y de Estudios Avanzados del IPN, Irapuato 36824, Mexico

* Correspondence: iaguilar@inmegen.gob.mx (I.A.-O.); enrique.morett@ibt.unam.mx (E.M.)

Abstract: Indigenous people from the American continent, or Native Americans, are underrepresented in the collective genomic knowledge. A minimal percentage of individuals in international databases belong to these important minority groups. Yet, the study of native American genomics is a growing field. In this work, we reviewed 56 scientific publications where ancient or contemporary DNA of Native Americans across the continent was studied by array, whole-exome, or whole-genome technologies. In total, 13,706 native Americans have been studied with genomic technologies, of which 1292 provided whole genome samples. Data availability is lacking, with barely 3.6% of the contemporary samples clearly accessible for further studies; in striking contrast, 96.3% of the ancient samples are publicly available. We compiled census data on the home countries and found that 607 indigenous groups are still missing representation in genomic datasets. By analyzing authorship of the published works, we found that there is a need for more involvement of the home countries as leads in indigenous genomic studies. We provide this review to aid in the design of future studies that aim to reduce the missing diversity of indigenous Americans.

Keywords: Native American; whole-genome sequencing; population genomics; data availability



Citation: Aguilar-Ordoñez, I.; Guzmán-Linares, J.; Ballesteros-Villascán, J.; Mirón-Toruño, F.; Pérez-González, A.; García-López, J.; Cruz-López, F.; Morett, E. A Tale of Native American Whole-Genome Sequencing and Other Technologies. *Diversity* **2022**, *14*, 647. <https://doi.org/10.3390/d14080647>

Academic Editor: Michael Wink

Received: 25 March 2022

Accepted: 14 May 2022

Published: 12 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In 2021, humanity celebrated 20 years since the publication of the first draft of the human genome [1], a document that gave rise to an era of self-exploration for our species at a molecular level. That human genome draft, thoroughly refined up to its most recent version [2], is still being used as the reference against which we compare genomic sequences from other individuals to detect genomic variations. However, one single genomic reference will not be enough to understand the whole human species; therefore, the complete sequences of a large number of individuals from different ancestries are required to better represent the human genomic variability [3].

The promise of genomic medicine, understood as the ability to provide personalized diagnosis and treatment for each patient, is based on an exhaustive study of human genomic variation. Genomic technologies such as array genotyping, exome sequencing, and whole-genome sequencing have allowed the study of this variation at a global scale [4,5], mainly by finding and cataloging single-nucleotide polymorphisms (SNPs) and more complex structural variations (SVs) [6] on DNA. However, most of the studies have focused on describing population genomics in high-income countries, leaving a gap in the potential understanding of the genomics underlying health and disease processes in the rest of the world. The underrepresentation of non-European populations in genomic science has been well documented and is a current topic of discussion [7–10]. Less than 1% of individuals

included in Genome-Wide Studies (or GWAS, which are studies aimed to find statistical relationships between traits and genomic variation) have indigenous or Native American (NatAm) ancestry [10]. Genomic studies of underrepresented people provide potential benefits for these populations, i.e., healthcare applications and policy making, or even anthropological information derived from genomics. For example, a genomic study found a genetic variant associated with higher risk of arrhythmia and sudden cardiac death in the Gitksan First Nation of Canada, followed by improved diagnosis and medical care for those affected [11]; in New Zealand, a study including the Ngāti Porou tribe identified associations between variants and gout disease, providing scientific evidence to improve diagnosis and treatment [11]; and in Mexico, genomic studies helped to promote the official recognition of the Afro-Mexican people, as evidenced by the inclusion of the option to self-identify as “Afro-Mexican” in the national population census conducted periodically by the government [12].

Latin America (LatAm) is home to diverse groups of genomically underrepresented NatAm ancestries (Figure 1), but LatAm population genomic studies have been sparse. This may be due to the fact that many of the LatAm countries face low funding for ambitious projects, and limited technical capabilities for locally analyzing genomic data (i.e., human resources and access to high-throughput hardware or cloud computing required for massive data processing). Still, LatAm has been the source for a modest amount of population genomic data generated either by local projects or as part of international efforts with global scopes.

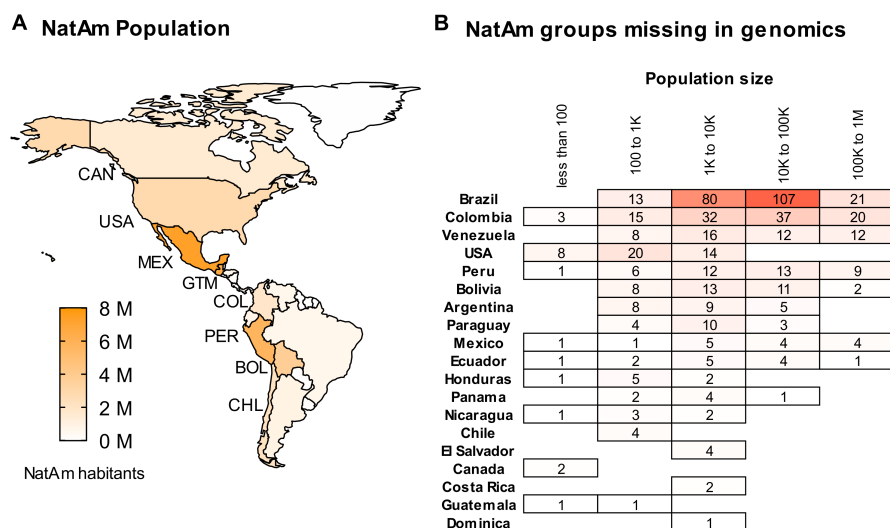


Figure 1. Native American home countries and the missing diversity. **(A)** The map highlights the indigenous population distribution across the continent, with labels on the countries with NatAm populations of over one million (Canada, USA, Mexico, Guatemala, Colombia, Peru, Bolivia, and Chile); color in the map indicates the number of Native American inhabitants in each country. **(B)** The heatmap shows the number of indigenous groups that do not have representation in genomic databases; these were found by cross-referencing census data with ethnicities reported in the published genomic studies included in this review.

Even if genomic data from NatAms are scarce and scattered, they exist, and genomic researchers (especially those in budget-constrained LatAm countries) can leverage one important property of data: when gathered, they grow. Thus, it is important to have a compendium of every research effort for genotyping or sequencing NatAm populations thus far. This information can become a solid reference for researchers aiming to complete the genomic landscape of underrepresented groups, for example by letting the researcher know which of the NatAm groups are already represented in population genomics, and which have very few or no individuals studied at all.

In this review, we compiled published studies for high-throughput genotyping and/or sequencing projects that included individuals reported as representatives of NatAms. We

classified the studies according to the genomic technology used (array, exome, or whole genome) and the data availability of each project (public or private). The NatAm studies under review included contemporary (or modern) individuals identified as members or descendants of particular ethnic groups, and/or ancient remains of individuals related to NatAm. We summarized the NatAm groups that have been included in the compiled studies, allowing us to determine which other groups are still in need of representation (Figure 1).

2. A Summary of Native American Populations

Evidence suggests that ancestors entered America by different routes and at different times during the peopling of the continent [13–17]. The most accepted theory of NatAm settlement began in the late Pleistocene with the migration of human ancestors from north-eastern Asia into the American continent through the Bering land bridge approximately ~23 ka (thousand years ago) [18]. These first Americans spread widely through the continent following a north to south axis [15,18], settling into different regions and groups. Over the course of millennia, the descendants of these first travelers experienced various degrees of genetic and cultural admixture, divergence, isolation, and environmental adaptations, giving rise to the ancient cultures of the American continent [14,15,19]. It is worth clarifying that the NatAm term we use in this review is defined as an individual who inhabited the American continent before the arrival of European conquerors, but also an individual belonging to a modern native/indigenous community [20] descendant of pre-conquest native groups.

Currently, the American continent can be classified in two great regions: North America and Latin America and the Caribbean, with 55 countries distributed across these regions [21]. These countries are home to diverse NatAm groups. Official census information about these groups may be up to date or scarce depending on the home country. We compiled government and third-party census information about contemporary NatAm groups and found that there are an estimated 37,036,134 contemporaneous NatAm people (Supplementary Table S1) [22–48]. Countries with the highest NatAm populations are Mexico (~7.3 million), Guatemala (~6.5 million), Peru (~5.7 million), and Bolivia (~4.1 million). Meanwhile, regarding group diversity, the USA, Brazil, and Colombia are home to more than 100 different NatAm groups each. Many of the NatAm groups remain unrepresented, or missing, in modern genomics (Figure 1).

3. The Technology behind Population Genomics

The cost of DNA genotyping and sequencing have decreased thanks to the commercial success and establishment of next-generation sequencing (NGS), thus changing the ways of studying genetic diversity and human diseases [49]. Let us summarize some of the technical and practical characteristics of the different technologies used for digitizing a human genome, namely, whole-exome and whole-genome sequencing, and array genotyping.

Whole-exome sequencing (WES) provides sequence bits of the coding regions of the genome. Although the coding region covers barely 1.5% of the genome, it includes up to 85% of the genomic variants related to Mendelian diseases [50]. The fact that WES can cover such a functional part of the genome can be leveraged for population genetics to design studies focused on specific phenotypes, or to survey adaptation and selection hypotheses [51]. Another technology known as whole-genome sequencing (WGS) provides the sequence of the whole ~3 billion base pairs in a human genome (in contrast to the region-constrained WES). The difference in sequencing yield (the “amount” of genomes being sequenced) between WES and WGS also comes with a difference in price. In general, given the same research budget, WES will be able to sequence more samples than WGS, but WGS would cover the complete genome of each sample. Thus, the exome option may allow us to ask specific genomic questions in more samples, while the whole genome captures the complete picture, enabling the future broader use of data for times when we know more about variation outside the reaches of the coding genome, or to analyze complex genomic

rearrangements [52–54]. WGS also enables the analysis of otherwise overlooked regions of the genome. For example, inferring the impact of variation in regulatory regions, such as those identified by ENCODE [55], is still a complicated issue. Since there is not much knowledge about regulatory variation (in contrast with coding variation), this becomes a cycle of the unknown: WGS may not be performed because not much is known about those extra-coding regions, and not much is known about those regions because it is mostly unexplored.

There is a third commonly used technique to genetically survey populations called DNA array genotyping. DNA arrays are collections of probes anchored to a solid support that test for the existence of thousands to millions of specific genetic variants in the sample being probed [56]. If WGS sees the whole genome picture, and WES looks at ~1% of it, then DNA arrays focus on very specific pixels in the picture. These key points usually have proven importance as markers for ancestry, phenotype–genotype associations, or biomarkers in general. DNA arrays can be considered low-cost in comparison to WES and WGS, but one must keep in mind that arrays can only efficiently look at already known interesting variants. DNA arrays are blind to novel and unknown variation. For the study of populations previously underrepresented in genomic studies, using DNA array technology for genotyping might not be the best strategy, since important variation may occur in some populations while not in others [57].

DNA array, WES, and WGS are not mutually exclusive strategies. In fact, it is common in population genomics to combine them as steps in large projects, or to use their specific advantages as complementary information. Pilot studies with DNA arrays can be used to assess the inclusion criteria for the more expensive WES or WGS [58]. Meanwhile, WES and WGS could explore regional variation to reveal new markers that could improve DNA array surveying of subsequent projects focused on specific phenotype–genotype relationships [59]. This synergy between technologies can improve the collective knowledge required, for example, to improve imputation strategies in genome-wide association studies in specific populations [60,61].

Every one of the techniques briefly summarized has its pros and cons, but there seems to be an increasing trend to use WGS for population sequencing projects. The promise of personalized medicine pushes for the implementation of personal WGS as common clinical practice in economies that can afford it [62], and there is clear acknowledgement of the importance of collective whole-genome information in modern medicine and molecular anthropology [63–68]. For all of the above, it is likely that WGS will be established as the gold standard for NatAm population genomics in the near future.

4. To Sequence or Not to Sequence?

The genome in every NatAm is an important vessel for history. Native genomes tell ancient stories but also enlighten the future of genomic science. Comparing genomes from different native groups can help scientists infer the patterns of migration and other demographic milestones that happened thousands of years ago during the peopling of the continent [69]. Meanwhile, analyzing the presence of currently known biomarkers [59] could potentially aid in future recommendations for better treatments and/or healthcare policy making to improve medical services in the otherwise marginalized NatAm groups.

To access the biological information within the genome, we need to digitize it via some of the three aforementioned genotyping technologies (array genotyping, WES, or WGS). The first NatAm population genomic projects used array technology. Many years later, although array technology is still a viable alternative, NGS is increasingly being used in population sequencing projects. Since the differences between WES and WGS can be simplified to scope and price, when population sequencing projects are designed, one of the decisions that must be made is whether to sequence more samples by exome sequencing, or fewer samples but with whole genome coverage. This is not a trivial question, more so in developing economies, including most of the countries that are home to NatAm groups. In these conditions, both answers are right, but only if the data produced are meant to be shared or integrated in larger datasets, because genomic data are cumulative.

Even a modest project sequencing 10 exomes, if executed right, can be an important source of genomic diversity in the great scheme of science because future projects can integrate those data with other sources. For example, increasing the number of publicly available genomes can be used to improve variant imputation (an important step for precise GWAS projects); simulations reported by Jiménez-Kauffman et al. [61] establish that at least 3000 NatAm whole genomes would be required to equal imputation accuracy to that of European descent populations. The increase in numbers should come with an increase in the diversity, and the study of diverse individuals can help scientists identify regional variation, such as the description of the D4h3 mitochondrial haplotype restricted to Pacific coastal populations [15]. Since more data potentially means more knowledge of NatAm genomics, molecular anthropology can also leverage cumulative genomic data to clarify population divergence times, migration routes, and even the origin of ancestral remains.

5. Sequencing, by the Numbers

From the first published project in 2009, the number of works in NatAm genomics has steadily increased, up to a total of 56 projects in early 2022 [4,61,70,71]. This is a good trend towards closing the gap for underrepresented groups in modern genomics. Of all the projects we reviewed, 26 included NatAm individuals in genotyping array-based studies, while 11 projects used WES and 29 used WGS (Supplementary Table S2) [4,18,19,61,70–121].

All three genomic technologies are still being used in recent years. There seems to be an increasing number of WGS projects being published; however, WES and array publications usually involve larger cohorts. Thus, the increase in the number and frequency of works comes along with a general increase in the number of studied individuals (Figure 2). To date, a total of 13,706 NatAm individuals have been studied with these genomic technologies: 4009 with arrays, 8405 with WES, and 1292 with WGS. Regarding NatAm WGS (the one that provides a broader picture of population genomics), the 1000 whole genomes studied mark was reached in 2018. This means there are over 3,200,000,000,000 surveyed nucleotides already published, painting the picture of the diversity and the story of the American continent. Since most of this biological information has been analyzed by independent (not directly related) projects, we insist that the long-term transcendence and importance of these data relies on their availability.

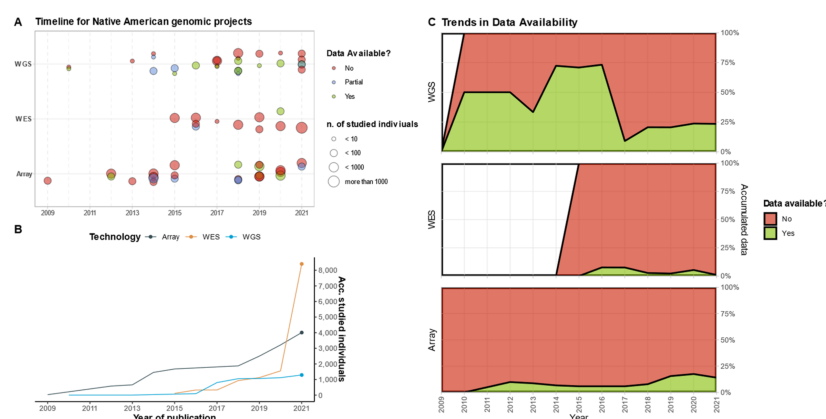


Figure 2. The history of NatAm genomics. (A) Projects across time. (B) Total increase in the number of NatAm individuals included in genomic studies. (C) Dynamics in NatAm data sharing. Only 1077 of the digital samples generated in NatAm genomic projects are readily available, with varying shares of availability by technology used (array, WES or WGS).

6. About Data Availability and the Missing NatAm Groups

NatAm genomics as a field struggles to provide an accessible source of data, as shown by our analysis of data availability in the reviewed works (Figure 2). We followed the provided links for data acquisition in the published manuscripts and tagged each sample as available or not available (Supplementary Table S3). An “available” sample provides direct

information or links to download sequencing files through established repositories such as the SRA (Sequence Read Archive), ENA (European Nucleotide Archive), or similar, while a “not available” sample does not provide a download link or direct and clear information as to how to access the data. Samples “available upon request” were marked as “not available”. We analyzed the accumulated share of studied individuals and found that as of 2021, only 14% of array samples are available; for WES, there is only 0.9%, and for WGS, the available share is 23.4% (Figure 2). To date, 489 of the studied individuals are ancient samples, and 13,217 are contemporary. Strikingly, 96.3% of the ancient samples provide directly available data, in contrast with only 3.6% of the modern samples (Figure 3).

Projects with data available upon request provide different types of agreements. Some provide signable premade forms while others provide only the email for contact. Most of the researchers listed as contacts for requesting access remain adscripted to the institution that led the project; this ensures a continuity in the responsibility of data keeping at both personal and institutional levels.

As better discussed elsewhere [122,123], data sharing upon request may not be the best strategy, since it hinders the growth of genomics as a research field in developing countries. “Available upon request” may become a vague term and a burden for both parties involved because data keepers have extra work in managing requests, and requesters may be left guessing what the exact terms and conditions are to avoid rejection. The caution behind availability upon request can be understood because most of the NatAm groups in these studies are marginalized people who must not be aggravated by improper use or interpretation of their genomic context [124]; however, communication and collaboration between experts (not only in genomics, but also in legal, cultural, and anthropological matters) could find a secure, respectful way to make these data open for future researchers. Controlled access sharing in genomic repositories provides a good opportunity for open access of de-identified data in organized automated databases [122]. For example, Jiménez-Kaufmann et al. [61] recently provided NatAm genomic data through the European Genome-Phenome Archive (EGA) repository, with a Data Access Committee keeping access permissions; the requester is asked to compromise that data will be used only for genetic studies in non-commercial endeavors. In 2016, Mallick et al. [104] published NatAm data mixing public genomes deposited at ENA and controlled access genomes safeguarded in EGA, showing that combining access methods is possible when some informed consents allow public sharing while others do not; in this case, the commitment letter also includes the compromise that the requester will not re-distribute data or post them publicly. We invite readers to learn more “about the rights and interests that Indigenous communities might have in genomic data” as well as to find “several principles for sharing genomic data derived from Indigenous communities” in the work of Hudson et al. [125]. For recommended guidelines to study ancient DNA, we suggest the recent work of Alpaslan-Roodenberg et al. [126].

In total, 178 NatAm groups (contemporary and ancient) have been represented in population genomics. Array technology has been used to survey 152 NatAm groups, WES 19 groups, and WGS 70 groups. The Pima, Maya, Inuit, Nahua, Rapa Nui, and Aymara are among the most studied contemporary NatAm populations, while the Chumash, Dorset, Aleutian, Pericues, Rapa Nui, and Kawéskar are among the most studied by ancient DNA (Figure 3). In this review, we collected census data from the different countries that are home to contemporary NatAm populations (Supplementary Table S4). By cross-referencing these data with the genomic studies published, we found that currently, 607 NatAm groups are still missing representation in the collective genomic archive (Figure 1, Supplementary Table S5). Groups such as Tojolabals, Lacandons, and Seri are known for their distinctive variation (such as high divergence shown by F_{ST} , and signals of isolation) [123], but there is only a handful of genomic data available upon request. As we have insisted before, WGS provides the complete genomic context of the studied population, and thus, future genomic endeavors could greatly benefit from analyzing whole genomes in the missing NatAm populations.

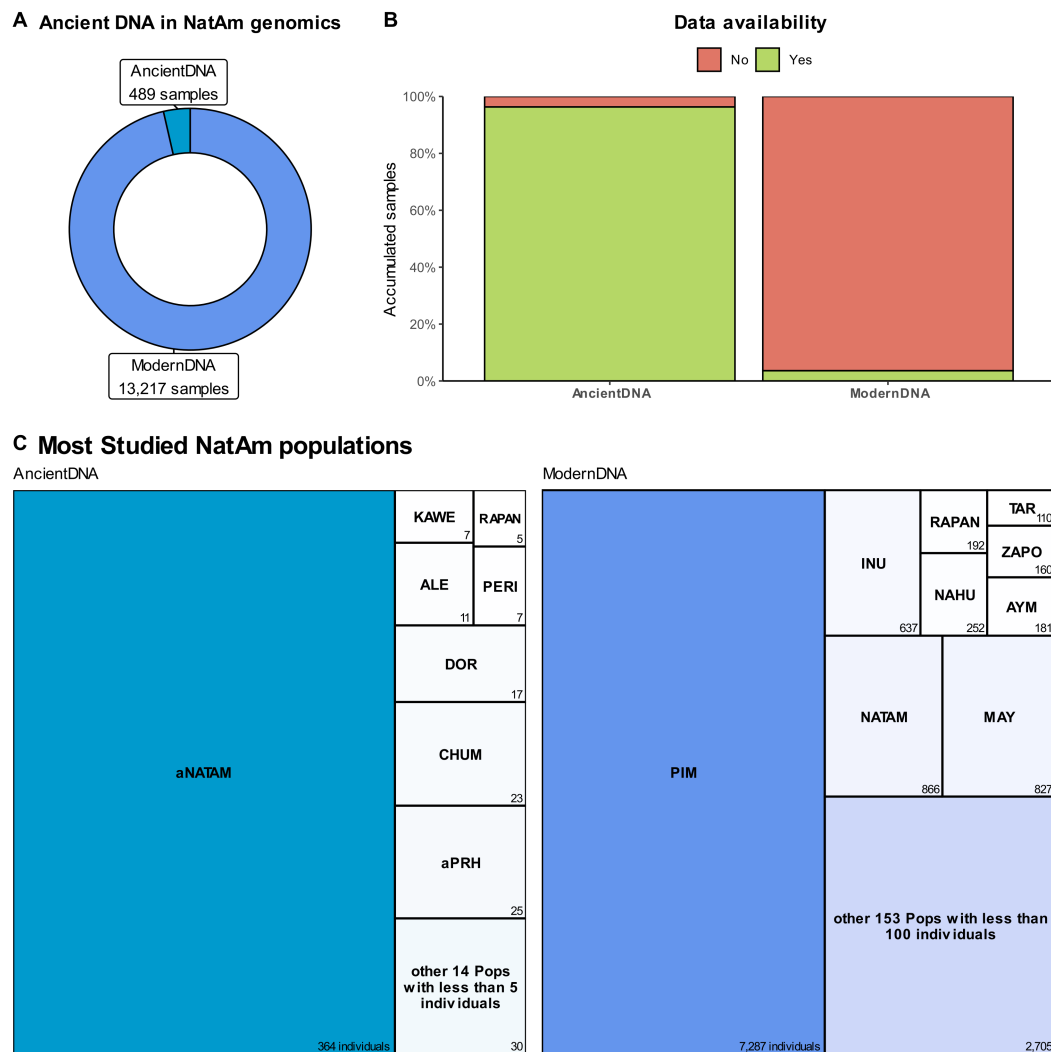


Figure 3. Overview of published NatAm genomic samples. (A,B) Ancient samples are a minimal fraction of the total NatAm genomic samples, yet they are mostly available, in contrast with contemporary samples. (C) Some NatAm groups have been thoroughly studied; Pima (PIM), Maya (MAY), and Inuit (INU) are some of the most studied contemporary groups, while Chumash (CHUM), Dorset (DOR), and Aleutian (ALE) are among the most represented in ancient datasets. The homologated codes and full name of the groups can be found in Supplementary Table S3.

In addition, we analyzed the country of origin for every research institution involved in the reviewed projects (Supplementary Table S6) [4,18,19,61,70–121]. The USA, Denmark, and Mexico led most of the projects (Figure 4). In the USA and Denmark, research groups that led the works are interested in past and present human diversity in general and have published several reports studying populations in different regions of the world, not only NatAm; the expertise and leadership from these countries is somewhat expected. The case for Mexico is interesting because in past years, the government and universities have invested in infrastructure and human capital for the development of regional genomics [127]; our brief analysis of authorships may be showing that Mexican research groups in population genomics have flourished as references in the field.

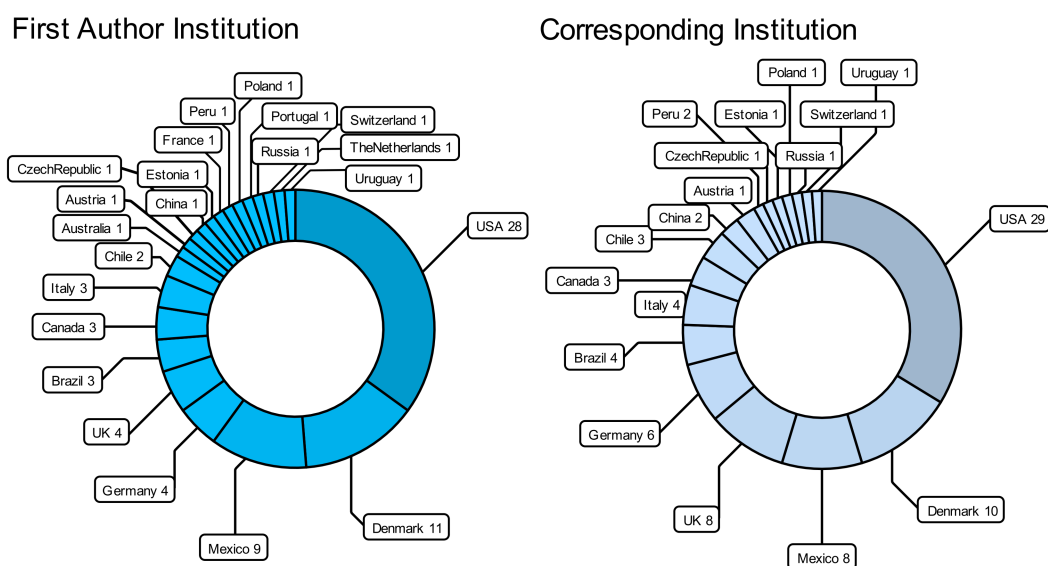


Figure 4. Country involvement in NatAm population genomic research. There is a lack of representation of NatAm home countries as project leads.

If we compare Figure 4 with Figure 1, it becomes evident that many of the home countries for the missing NatAm groups are also missing representation as leads in research projects. This may be because in developing economies, advances in genomics are, most of the time, perceived to be out of reach, both financially and logistically [128]. Yet, a question remains: how do we better involve NatAm home countries in future NatAm genomics?

The answer may be far from simple because successful population genomics initiatives must consider the political will and institutional leadership of each country [127,129]. Political support is crucial to acquire funding to develop research studies [130], but researchers must also make sure that political decision makers have realistic expectations of the benefits and the timespan of genomic projects [129]. Lead researchers must also plan strategies to raise public interest and, of utmost importance, to involve native communities in the project [129]. LatAm home countries as project leads can benefit from international collaborations by joining forces with more experienced developed countries, from which they can receive knowledge, access to infrastructure, training, and open channels to rapidly communicate findings [130].

7. Concluding Remarks

The study of NatAm genomics is a young field, but it has grown rapidly (Figure 2). Researchers must adapt equally as fast to the challenges behind the ethics of population genomic studies. The reasoning behind research projects must also include questions about how they are important for NatAm communities [131]. It is important that NatAm communities become actively involved in conversation and decision making about the treatment of the genomic data and the reaches of the research behind the projects that include them [132].

There are still hundreds of indigenous groups lacking genomic representation. Nineteen of these groups have reported populations of more than 100,000 people. If future population sequencing projects would focus on these groups, the potential impact of the genomic knowledge gained maybe would be far greater. Unfortunately, there is little openness for data sharing in contemporary NatAm genomics. This can be understood from an ethical point of view, since many of the indigenous groups in America still face marginalization and alienation in their home countries, such that protection of the genomic heritage and avoidance of future misuse of this information must always be a priority. Future NatAm projects must carefully acknowledge the risks of collecting, handling, and publishing data from indigenous populations, mainly to avoid repeating known mistakes

such as the lack of informed consent for secondary use of genetic data (as was the case for the Havasupai Tribe in Arizona, USA [133], and the Nuu-chah-nulth in British Columbia, Canada [134]), or the inaccurate association of debatable specific phenotypes (as in the negative representation in scientific publications of the “warrior gene” in the indigenous Maori people of New Zealand [135]).

However, given the possibility to anonymize genomic data, and the existence of public, well-established, and secure repositories such as SRA, ENA, and EGA, there might be options to clearly share data without compromising the genomic security of indigenous people. It is imperative to try and solve the issue of data sharing in NatAm genomics, since much can be gained by gathering the efforts behind the projects we reviewed. As recently shown [61], available data from multiple sources can be gathered to create a larger dataset that better represents the NatAm genomic context.

Our study focused on bibliometrics and informatics, but we did not review the informed consent behind every project. It may be worth deeply analyzing the reasons behind the wall in non-public NatAm datasets and to try to coordinate efforts with the respective corresponding authors to free and integrate the indigenous data in a collaborative effort, without compromising the security and consent of the indigenous participants. Lastly, there is a need for more involvement of LatAm countries in population genomic studies, not only as collaborators but as lead institutions. A more direct involvement of the home countries should also aim to improve the involvement of the indigenous communities, so communities and home countries risk not becoming merely sample providers, but active participants in the journey of collective genomic discovery. To aid in future projects aiming to close the gap in these underrepresented NatAm groups, we provide the complete list of missing groups as Supplementary Table S4.

8. Methods

8.1. Data Collection

We searched the PubMed search engine with the following terms. For WGS projects: (((Native American whole-genome sequencing) OR Native American Ancient DNA)) NOT (whole exome sequencing)) NOT (genotyping arrays). For WES projects: (((Native American whole exome sequencing) OR Native American Ancient DNA)) NOT (whole-genome sequencing)) NOT (genotyping arrays). For array projects: (((Native American genotyping arrays) OR Native American Ancient DNA genotyping arrays)) NOT (whole-genome sequencing)) NOT (whole exome sequencing). Then, we used ResearchRabbit (<https://www.researchrabbit.ai/>, accessed on 19 February 2022) to find papers similar to our initial findings. Final inclusion criteria for papers included in the review are works that used WGS, WES, or array analysis in NatAm; works including contemporary or ancient NatAm samples; and works where data availability is reported.

8.2. Census Information on NatAm Home Countries

Census information on NatAm populations was sought for each home country (Supplementary Table S4 provides all the sources and links used). When no official census could be found, we used data from The Indigenous World 2021 [136].

8.3. Data Handling and Visualization

Data wrangling and visualization were performed in R, with several useful packages [137–152]. The code to reproduce every analysis and raw figures can be found as a R markdown script in the following git repository: <https://github.com/laguilaror/natam-review>, accessed on 24 March 2022, it includes a copy of the Supplementary Tables in spreadsheet format.

Supplementary Materials: All supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/d14080647/s1>. Table S1: Native American Groups and Estimated Population per home country (accessed on 1 March 2022). Table S2: Full list of Native American population

genomics projects. Table S3: Full list of Native American samples in population genomics. Table S4: Full list of Native American groups and populations (accessed on 1 March 2022). Table S5: Missing Native American groups in population genomics (accessed on 1 March 2022). Table S6: Authorship information on Native American genomic projects. References [22–48] cited in the Table S1, references [4,18,61,70–121] cited in the Table S2, references [117–121] cited in the Table S6 of the Supplementary Materials.

Author Contributions: Study design: I.A.-O. and E.M. Data collection: J.G.-L. (Josué Guzmán-Linares) and F.C.-L. Data analysis: I.A.-O. and J.G.-L. (Josué Guzmán-Linares). Writing: I.A.-O., J.G.-L. (Josué Guzmán-Linares), E.M., J.B.-V., F.M.-T., A.P.-G. and J.G.-L. (José García-López). Funding: I.A.-O. and E.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by CONACYT PhD program (I. A-O: 480576), CONACYT SNI research assistant program (J. G-L: 19618, F. M-T: 1076403, AP. P-G.: 904078, JE. G-L: 1076378), and BUAP through the “Programa Interinstitucional para el Fortalecimiento de la Investigación y el Posgrado del Pacífico DELFIN” (J. G-L: 04719).

Institutional Review Board Statement: This work compiles bibliometrics and anonymized metadata directly available from the original reports regarding human genomics research; no new original genomic data was produced. This review only includes projects that were conducted in accordance with the Declaration of Helsinki; particular Ethics Committee information can be found in each original report.

Data Availability Statement: The code to reproduce analyses and raw figures can be found as an R markdown script in the following git repository: <https://github.com/Iaguilaror/natam-review>, accessed on 24 March 2022, it includes a copy of the Supplementary Tables in spreadsheet format.

Acknowledgments: The authors would like to thank Zaira Sanchez-Cuapio, Emmanuel Salazar, and Nidia Campusano for helping curate the data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [PubMed]
2. Nurk, S.; Koren, S.; Rhie, A.; Rautiainen, M.; Bzikadze, A.V.; Mikheenko, A.; Vollger, M.R.; Altemose, N.; Uralsky, L.; Gershman, A.; et al. The complete sequence of a human genome. *Science* **2022**, *376*, 44–53. [CrossRef] [PubMed]
3. Ballouz, S.; Dobin, A.; Gillis, J.A. Is it time to change the reference genome? *Genome Biol.* **2019**, *20*, 159. [CrossRef] [PubMed]
4. Bergström, A.; McCarthy, S.A.; Hui, R.; Almarri, M.A.; Ayub, Q.; Danecek, P.; Chen, Y.; Felkel, S.; Hallast, P.; Kamm, J.; et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* **2020**, *367*, eaay5012. [CrossRef]
5. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [CrossRef]
6. Collins, R.L.; Brand, H.; Karczewski, K.J.; Zhao, X.; Alfoldi, J.; Francioli, L.C.; Khera, A.V.; Lowther, C.; Gauthier, L.D.; Wang, H.; et al. A structural variation reference for medical and population genetics. *Nature* **2020**, *581*, 444–451. [CrossRef]
7. Need, A.C.; Goldstein, D.B. Next generation disparities in human genomics: Concerns and remedies. *Trends Genet.* **2009**, *25*, 489–494. [CrossRef]
8. Bustamante, C.D.; De La Vega, F.M.; Burchard, E.G. Genomics for the world. *Nature* **2011**, *475*, 163–165. [CrossRef]
9. Popejoy, A.B.; Fullerton, S.M. Genomics is failing on diversity. *Nature* **2016**, *538*, 161–164. [CrossRef]
10. Sirugo, G.; Williams, S.M.; Tishkoff, S.A. The Missing Diversity in Human Genetic Studies. *Cell* **2019**, *177*, 26–31. [CrossRef]
11. Caron, N.R.; Chongo, M.; Hudson, M.; Arbour, L.; Wasserman, W.W.; Robertson, S.; Correard, S.; Wilcox, P. Indigenous Genomic Databases: Pragmatic Considerations and Cultural Contexts. *Front. Public Health* **2020**, *8*, 11. [CrossRef]
12. Guglielmi, G. Facing up to injustice in genome science. *Nature* **2019**, *568*, 290–293. [CrossRef]
13. Pedersen, M.W.; Ruter, A.; Schweger, C.; Friebe, H.; Staff, R.A.; Kjeldsen, K.K.; Mendoza, M.L.Z.; Beaudoin, A.B.; Zutter, C.; Larsen, N.K.; et al. Postglacial viability and colonization in North America’s ice-free corridor. *Nature* **2016**, *537*, 45–49. [CrossRef]
14. Willerslev, E.; Meltzer, D.J. Peopling of the Americas as inferred from ancient genomics. *Nature* **2021**, *594*, 356–364. [CrossRef]
15. Skoglund, P.; Reich, D. A genomic view of the peopling of the Americas. *Curr. Opin. Genet. Dev.* **2016**, *41*, 27–35. [CrossRef]
16. Potter, B.A.; Baichtal, J.F.; Beaudoin, A.B.; Fehren-Schmitz, L.; Haynes, C.V.; Holliday, V.T.; Holmes, C.E.; Ives, J.W.; Kelly, R.L.; Llamas, B.; et al. Current evidence allows multiple models for the peopling of the Americas. *Sci. Adv.* **2018**, *4*, eaat5473. [CrossRef]
17. Bennett, M.R.; Bustos, D.; Pigati, J.S.; Springer, K.B.; Urban, T.M.; Holliday, V.T.; Reynolds, S.C.; Budka, M.; Honke, J.S.; Hudson, A.M.; et al. Evidence of humans in North America during the Last Glacial Maximum. *Science* **2021**, *373*, 1528–1531. [CrossRef]

18. Raghavan, M.; Steinrücken, M.; Harris, K.; Schiffels, S.; Rasmussen, S.; DeGiorgio, M.; Albrechtsen, A.; Valdiosera, C.; Ávila-Arcos, M.C.; Malaspina, A.; et al. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* **2015**, *349*, aab3884. [CrossRef]
19. Scheib, C.L.; Li, H.; Desai, T.; Link, V.; Kendall, C.; Dewar, G.; Griffith, P.W.; Morseburg, A.; Johnson, J.R.; Potter, A.; et al. Ancient human parallel lineages within North America contributed to a coastal expansion. *Science* **2018**, *360*, 1024–1027. [CrossRef]
20. Bolnick, D.A.; Raff, J.A.; Springs, L.C.; Reynolds, A.W.; Miró-Herrans, A.T. Native American Genomics and Population Histories. *Annu. Rev. Anthropol.* **2016**, *45*, 319–340. [CrossRef]
21. United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2019*; UN: New York, NY, USA, 2019; Volume II.
22. Instituto Nacional de Estadística y Geografía. Censo de Población y Vivienda 2020. 2021. Available online: <https://www.inegi.org.mx/programas/ccpv/2020/#Tabulados> (accessed on 8 March 2022).
23. Instituto Nacional de Estadística. Portal de Resultados del Censo 2018. Available online: <https://www.censopoblacion.gt/explorador> (accessed on 8 March 2022).
24. Base de Datos de Pueblos Indígenas u Originarios. Lista de Pueblos Indígenas u Originarios. Available online: <https://bdpi.cultura.gob.pe/pueblos-indigenas> (accessed on 8 March 2022).
25. Instituto Nacional de Estadística. Censos. INE. Available online: <https://www.ine.gob.bo/index.php/censos-y-banco-de-datos/censos/> (accessed on 8 March 2022).
26. Norris, T.; Vines, P.L.; Hoeffel, E.M. *The American Indian and Alaska Native Population: 2010*; US Department of Commerce, Economics and Statistics Administration, US Census Bureau: Suitland, MD, USA, 2010; Volume 21.
27. Censo 2017. 2017. Available online: <https://www.censo2017.cl/descargas/home/sintesis-de-resultados-censo2017.pdf> (accessed on 17 May 2022).
28. Departamento Administrativo Nacional de Estadística. Grupos Étnicos Información Técnica. 2018. Available online: <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/grupos-etnicos/informacion-tecnica> (accessed on 8 March 2022).
29. Government of Canada SC. Aboriginal Population Profile, 2016 Census—Canada [Country]. 2018. Available online: <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/abpopprof/details/page.cfm?Lang=E&Geo1=PR&Code1=01&Data=Count&SearchText=Canada&SearchType=Begins&B1=Aboriginal%20peoples&C1=All&SEXID=1&AGEID=1&RESGEOID=1> (accessed on 8 March 2022).
30. INDEC (Instituto Nacional de Estadística y Censos de la República Argentina). Pueblos Originarios. 2010. Available online: <https://www.indec.gob.ar/indec/web/Nivel4-Tema-2--21--99> (accessed on 8 March 2022).
31. Silverio Chisaguano, M. *La Población Indígena del Ecuador*; Instituto Nacional de Estadística y Censos (INEC): Loja, Ecuador, 2006; p. 39.
32. Gerencia General de Estadísticas Demográficas. *Resultados Población Indígena. XIV Censo de Población y Vivienda 2011*; Instituto Nacional de Estadística: Caracas, Venezuela, 2011; p. 42. Available online: <http://www.ine.gov.ve/documentos/Demografia/CensodePoblacionyVivienda/pdf/ResultadosBasicos.pdf> (accessed on 9 March 2022).
33. Quadro Geral dos Povos—Povos Indígenas No Brasil. 2021. Available online: https://piib.socioambiental.org/pt/Quadro_Geral_dos_Povos (accessed on 9 March 2022).
34. Instituto Nacional de Estadística de Honduras. Tomo 6. Grupos Poblacionales. 2015. Available online: https://www.ine.gob.hn/publicaciones/Censo_2013/06Tomo-VI-Grupos-Poblacionales/cuadros.html (accessed on 9 March 2022).
35. Davis, E. *Diagnóstico de la Población Indígena de Panamá con Base en los Censos de Población y Vivienda de 2010*; Instituto Nacional de Estadística y Censo (INEC): Panama City, Panama, 2010; p. 186. Available online: https://inec.gob.pa/archivos/P6571INDIGENA_FINAL_FINAL.pdf (accessed on 9 March 2022).
36. Oficina Pueblos Indígenas. Nota técnica de país sobre cuestiones de los pueblos indígenas. República de Nicaragua: Centro para la Autonomía y Desarrollo de los Pueblos Indígenas (CADPI). 2017. Available online: https://www.ifad.org/documents/38714170/40258424/nicaragua_ctn_s.pdf (accessed on 19 May 2022).
37. Otazú, N.; Bazán, R.; Leiva, V.M. *Censo de Comunidades de los Pueblos Indígenas Resultados Finales 2012*; Dirección General de Estadística, Encuestas y Censos (DGEEC): Asuncion, Paraguay, 2012; p. 140. Available online: https://www.ine.gov.py/Publicaciones/Biblioteca/documento/c3c9_Censo%20de%20Comunidades%20de%20los%20Pueblos%20Indigenas%20Resultados%20Finales%202012.pdf (accessed on 9 March 2022).
38. Instituto Nacional de Estadística y Censos. Costa Rica: Población Indígena por Pertenencia a un Pueblo Indígena, Según Provincia y Sexo. 2011. Available online: <https://www.inec.cr/sites/default/files/documentos/poblacion/estadisticas/resultados/resocialcenso2011--03.xls.xls> (accessed on 22 March 2022).
39. International Work Group for Indigenous Affairs (IWGIA). Indigenous Peoples in Guyana. Available online: <https://iwgia.org/en/guyana.html> (accessed on 9 March 2022).
40. Rodríguez, P.; Justo, C.; Miguel, C.; Olivera, J.; Martino, D. *Población Indígena en Uruguay y su Vínculo con el Bosque. Proyecto REDD+ Uruguay*; Ministerio de Ganadería, Agricultura y Pesca-Ministerio de Vivienda, Ordenamiento Territorial y Medio Ambiente: Montevideo, Uruguay, 2020; p. 47. Available online: https://www.gub.uy/ministerio-ganaderia-agricultura-pesca/sites/ministerio-ganaderia-agricultura-pesca/files/documentos/publicaciones/1.%20Informe_PI_y_BN.pdf (accessed on 9 March 2022).
41. Grønlands Statistik. Available online: <https://stat.gl/dialog/topmain.asp?lang=da&subject=Befolkning&sc=BE> (accessed on 9 March 2022).

42. The Statistical Institute of Belize. *Belize Population and Housing Census (Country Report 2010)*; The Statistical Institute of Belize: Belmopan, Belize, 2013; p. 176. Available online: http://sib.org.bz/wp-content/uploads/2010_Census_Report.pdf (accessed on 20 May 2022).
43. General Bureau of Statistics. *Mosaic of the Surinamese People*. 2019. Available online: <https://statistics-suriname.org/en/mosaic-of-the-surinamese-people/> (accessed on 20 May 2022).
44. Censo de Población y Vivienda 2007 (Población). El Salvador: Dirección General de Estadística y Censos (DIGESTYC). 2007. Available online: <http://www.digestyc.gob.sv/index.php/temas/des/poblacion-y-estadisticas-demograficas/censo-depoblacion-y-vivienda/poblacion-censos.html> (accessed on 9 March 2022).
45. International Work Group for Indigenous Affairs (IWGIA). *Indigenous Peoples in French Guiana*. Available online: <https://iwgia.org/en/french-guiana.html> (accessed on 9 March 2022).
46. Population and Housing Census 2012. St. Vincent and the Grenadines: Statistical Office, Ministry of Finance, Government of St Vincent and the Grenadines. 2012. Available online: <https://redatam.org/binsvg/RpWebEngine.exe/Portal?BASE=SVG2012> (accessed on 20 May 2022).
47. Ethnic Groups by Sex 1991 2001 and 2011. Commonwealth of Dominica: Central Statistics Office of Dominica. Available online: <https://stats.gov.dm/subjects/demographic-statistics/ethnic-groups-by-sex-1991-2001-and-2011/> (accessed on 9 March 2022).
48. Government of St. Lucia. *The 2010 Population and Housing Census of St. Lucia*; The Central Statistics Office, Government of St. Lucia: Castries, Saint Lucia, 2010. Available online: <https://redatam.org/binlca/RpWebEngine.exe/Portal?BASE=PHC2010C&lang=ENG> (accessed on 9 March 2022).
49. Stephens, Z.D.; Lee, S.Y.; Faghri, F.; Campbell, R.H.; Zhai, C.; Efron, M.J.; Iyer, R.; Schatz, M.C.; Sinha, S.; Robinson, G.E. Big Data: Astronomical or Genomical? *PLoS Biol.* **2015**, *13*, e1002195. [[CrossRef](#)]
50. Ng, S.B.; Turner, E.H.; Robertson, P.D.; Flygare, S.D.; Bigham, A.W.; Lee, C.; Shaffer, T.; Wong, M.; Bhattacharjee, A.; Eichler, E.E.; et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **2009**, *461*, 272–276. [[CrossRef](#)]
51. Udupa, N.; Ronen, R.; Zhou, D.; Liang, J.; Stobdan, T.; Appenzeller, O.; Yin, Y.; Du, Y.; Guo, L.; Cao, R.; et al. Whole genome sequencing of Ethiopian highlanders reveals conserved hypoxia tolerance genes. *Genome Biol.* **2014**, *15*, R36. [[CrossRef](#)]
52. Francioli, L.C.; Menelaou, A.; Pulit, S.L.; Van Dijk, F.; Palamara, P.F.; Elbers, C.C.; Neerincx, P.B.T.; Ye, K.; Gurtev, V.; The Genome of the Netherlands Consortium; et al. Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat. Genet.* **2014**, *46*, 818–825.
53. Gilissen, C.; Hehir-Kwa, J.Y.; Thung, D.T.; van de Vorst, M.; van Bon, B.W.M.; Willemsen, M.H.; Kwint, M.; Janssen, I.M.; Hoischen, A.; Schenck, A.; et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature* **2014**, *511*, 344–347. [[CrossRef](#)]
54. Sanders, S.J.; Murtha, M.T.; Gupta, A.R.; Murdoch, J.D.; Raubeson, M.J.; Willsey, A.J.; Ercan-Sencicek, A.G.; DiLullo, N.M.; Parikshak, N.N.; Stein, J.L.; et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **2012**, *485*, 237–241. [[CrossRef](#)]
55. ENCODE Project Consortium; Dunham, I.; Birney, E.; Lajoie, B.; Sanyal, A.; Dong, X.; Greven, M.; Lin, X.; Wang, J.; Whitfield, T.W.; et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.
56. Gresham, D.; Dunham, M.J.; Botstein, D. Comparing whole genomes using DNA microarrays. *Nat. Rev. Genet.* **2008**, *9*, 291–302. [[CrossRef](#)]
57. Foo, J.N.; Liu, J.J.; Tan, E.K. Whole-genome and whole-exome sequencing in neurological diseases. *Nat. Rev. Neurol.* **2012**, *8*, 508–517. [[CrossRef](#)]
58. Poulsen, J.B.; Lescai, F.; Grove, J.; Bækvad-Hansen, M.; Christiansen, M.; Hagen, C.M.; Maller, J.; Stevens, C.; Li, S.; Li, Q.; et al. High-Quality Exome Sequencing of Whole-Genome Amplified Neonatal Dried Blood Spot DNA. *PLoS ONE* **2016**, *11*, e0153253. [[CrossRef](#)]
59. Gonzalez-Covarrubias, V.; Morales-Franco, M.; Cruz-Correa, O.F.; Martínez-Hernández, A.; García-Ortiz, H.; Barajas-Olmos, F.; Genis-Mendoza, A.D.; Martínez-Magaña, J.J.; Nicolini, H.; Orozco, L.; et al. Variation in Actionable Pharmacogenetic Markers in Natives and Mestizos From Mexico. *Front. Pharmacol.* **2019**, *10*, 1169. [[CrossRef](#)] [[PubMed](#)]
60. Rustagi, N.; Zhou, A.; Watkins, W.S.; Gedvilaite, E.; Wang, S.; Ramesh, N.; Muzny, D.; Gibbs, R.A.; Jorde, L.B.; Yu, F.; et al. Extremely low-coverage whole genome sequencing in South Asians captures population genomics information. *BMC Genom.* **2017**, *18*, 396. [[CrossRef](#)] [[PubMed](#)]
61. Jiménez-Kaufmann, A.; Chong, A.Y.; Cortés, A.; Quinto-Cortés, C.D.; Fernandez-Valverde, S.L.; Ferreyra-Reyes, L.; Cruz-Hervet, L.P.; Medina-Muñoz, S.G.; Sohail, M.; Palma-Martinez, M.J.; et al. Imputation Performance in Latin American Populations: Improving Rare Variants Representation with the Inclusion of Native American Genomes. *Front. Genet.* **2022**, *12*, 719791. [[CrossRef](#)] [[PubMed](#)]
62. Marshall, C.R.; Bick, D.; Belmont, J.W.; Taylor, S.L.; Ashley, E.; Dimmock, D.; Jobanputra, V.; Kearney, H.M.; Kulkarni, S.; Rehm, H. The Medical Genome Initiative: Moving whole-genome sequencing for rare disease diagnosis to the clinic. *Genome Med.* **2020**, *12*, 48. [[CrossRef](#)]
63. Bick, D.; Fraser, P.C.; Gutzeit, M.F.; Harris, J.M.; Hambuch, T.M.; Helbling, D.C.; Jacob, H.J.; Kersten, J.N.; Leuthner, S.R.; May, T.; et al. Successful Application of Whole Genome Sequencing in a Medical Genetics Clinic. *J. Pediatr. Genet.* **2017**, *6*, 61–76.

64. Cortés-Ciriano, I.; Lee, J.J.K.; Xi, R.; Jain, D.; Jung, Y.L.; Yang, L.; Gordenin, D.; Klimczak, L.J.; Zhang, C.-Z.; Pellman, D.S.; et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nat. Genet.* **2020**, *52*, 331–341. [\[CrossRef\]](#)
65. Dong, Z.; Xie, W.; Chen, H.; Xu, J.; Wang, H.; Li, Y.; Wang, J.; Chen, F.; Choy, K.W.; Jiang, H. Copy-Number Variants Detection by Low-Pass Whole-Genome Sequencing. *Curr. Protoc. Hum. Genet.* **2017**, *94*, 8.17.1–8.17.16. [\[CrossRef\]](#)
66. Kosugi, S.; Momozawa, Y.; Liu, X.; Terao, C.; Kubo, M.; Kamatani, Y. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. *Genome Biol.* **2019**, *20*, 117. [\[CrossRef\]](#)
67. Lippert, C.; Sabatini, R.; Maher, M.C.; Kang, E.Y.; Lee, S.; Arikan, O.; Harley, A.; Bernal, A.; Garst, P.; Lavrenko, V.; et al. Identification of individuals by trait prediction using whole-genome sequencing data. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 10166–10171. [\[CrossRef\]](#)
68. Nakagawa, H.; Fujita, M. Whole genome sequencing analysis for cancer genomics and precision medicine. *Cancer Sci.* **2018**, *109*, 513–522. [\[CrossRef\]](#)
69. Nielsen, R.; Akey, J.M.; Jakobsson, M.; Pritchard, J.K.; Tishkoff, S.; Willerslev, E. Tracing the peopling of the world through genomics. *Nature* **2017**, *541*, 302–310. [\[CrossRef\]](#)
70. Aguilar-Ordoñez, I.; Pérez-Villatoro, F.; García-Ortiz, H.; Barajas-Olmos, F.; Ballesteros-Villascán, J.; González-Buenfil, R.; Fresno, C.; Garcíaarrubio, A.; Fernández-López, J.C.; Tovar, H.; et al. Whole genome variation in 27 Mexican indigenous populations, demographic and biomedical insights. *PLoS ONE* **2021**, *16*, e0249773. [\[CrossRef\]](#)
71. Bongers, J.L.; Nakatsuka, N.; O'Shea, C.; Harper, T.K.; Tantaleán, H.; Stanish, C.; Fehren-Schmitz, L. Integration of ancient DNA with transdisciplinary dataset finds strong support for Inca resettlement in the south Peruvian coast. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 18359–18368. [\[CrossRef\]](#)
72. Day, S.E.; Muller, Y.L.; Koroglu, C.; Kobes, S.; Wiedrich, K.; Mahkee, D.; Baier, L.J. Exome Sequencing of 21 Bardet-Biedl Syndrome (BBS) Genes to Identify Obesity Variants in 6,851 American Indians. *Obesity* **2021**, *29*, 748–754. [\[CrossRef\]](#)
73. Spangenberg, L.; Fariello, M.I.; Arce, D.; Illanes, G.; Greif, G.; Shin, J.-Y.; Yoo, S.-K.; Seo, J.-S.; Robello, C.; Kim, C.; et al. Indigenous Ancestry and Admixture in the Uruguayan Population. *Front. Genet.* **2021**, *12*, 1818. [\[CrossRef\]](#)
74. García-Ortiz, H.; Barajas-Olmos, F.; Contreras-Cubas, C.; Cid-Soto, M.Á.; Córdova, E.J.; Centeno-Cruz, F.; Mendoza-Caamal, E.; Cicerón-Arellano, I.; Flores-Huacuja, M.; Baca, P.; et al. The genomic landscape of Mexican Indigenous populations brings insights into the peopling of the Americas. *Nat. Commun.* **2021**, *12*, 5942. [\[CrossRef\]](#)
75. Popović, D.; Molak, M.; Ziolkowski, M.; Vranich, A.; Sobczyk, M.; Vidaurre, D.U.; Agresti, G.; Skrzypczak, M.; Ginalski, K.; Lamnidis, T.C.; et al. Ancient genomes reveal long-range influence of the pre-Columbian culture and site of Tiwanaku. *Sci. Adv.* **2021**, *7*, eabg7261. [\[CrossRef\]](#)
76. Capodiferro, M.R.; Aram, B.; Raveane, A.; Migliore, N.R.; Colombo, G.; Ongaro, L.; Rivera, J.; Mendizábal, T.; Hernández-Mora, I.; Tribaldos, M.; et al. Archaeogenomic distinctiveness of the Isthmo-Colombian area. *Cell* **2021**, *184*, 1706–1723.e24. [\[CrossRef\]](#)
77. Ribeiro-Dos-Santos, A.M.; Vidal, A.F.; Vinasco-Sandoval, T.; Guerreiro, J.; Santos, S.; de Souza, S.J.; Ribeiro-dos-Santos, A. Exome Sequencing of Native Populations From the Amazon Reveals Patterns on the Peopling of South America. *Front. Genet.* **2020**, *11*, 548507. [\[CrossRef\]](#)
78. Piaggi, P.; Köroğlu, Ç.; Nair, A.K.; Sutherland, J.; Muller, Y.L.; Kumar, P.; Hsueh, W.C.; Kobes, S.; Shuldiner, A.R.; Kim, H.I.; et al. Exome Sequencing Identifies a Nonsense Variant in DAO Associated with Reduced Energy Expenditure in American Indians. *J. Clin. Endocrinol. Metab.* **2020**, *105*, e3989–e4000. [\[CrossRef\]](#)
79. Ioannidis, A.G.; Blanco-Portillo, J.; Sandoval, K.; Hagelberg, E.; Miquel-Poblete, J.F.; Moreno-Mayar, J.V.; Rodríguez-Rodríguez, J.E.; Quinto-Cortés, C.D.; Auckland, K.; Parks, T.; et al. Native American gene flow into Polynesia predating Easter Island settlement. *Nature* **2020**, *583*, 572–577. [\[CrossRef\]](#)
80. Fernandes, D.M.; Sirak, K.A.; Ringbauer, H.; Sedig, J.; Rohland, N.; Cheronet, O.; Mah, M.; Mallick, S.; Olalde, I.; Cullen, B.J.; et al. A genetic history of the pre-contact Caribbean. *Nature* **2021**, *590*, 103–110. [\[CrossRef\]](#)
81. Borda, V.; Alvim, I.; Mendes, M.; Silva-Carvalho, C.; Soares-Souza, G.B.; Leal, T.P.; Scliar, M.O.; Zamudio, R.; Zolani, C.; Padilla, C.; et al. The genetic structure and adaptation of Andean highlanders and Amazonians are influenced by the interplay between geography and culture. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 32557–32565. [\[CrossRef\]](#)
82. Ávila-Arcos, M.C.; McManus, K.F.; Sandoval, K.; Rodríguez-Rodríguez, J.E.; Villa-Islas, V.; Martín, A.R.; Luisi, P.; Peñaloza-Espinosa, R.I.; Eng, C.; Huntsman, S.; et al. Population History and Gene Divergence in Native Mexicans Inferred from 76 Human Exomes. *Mol. Biol. Evol.* **2020**, *37*, 994–1006. [\[CrossRef\]](#)
83. Zhou, S.; Xie, P.; Quibion, A.; Ambalavanan, A.; Dionne-Laporte, A.; Spiegelman, D.; Bourassa, C.V.; Xiong, L.; Dion, P.A.; Rouleau, G.A. Genetic architecture and adaptations of Nunavik Inuit. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 16012–16017. [\[CrossRef\]](#)
84. Vidal, E.A.; Moyano, T.C.; Bustos, B.I.; Pérez-Palma, E.; Moraga, C.; Riveras, E.; Montecinos, A.; Azócar, L.; Soto, D.C.; Vidal, M.; et al. Whole Genome Sequence, Variant Discovery and Annotation in Mapuche-Huilliche Native South Americans. *Sci. Rep.* **2019**, *9*, 2132. [\[CrossRef\]](#)
85. Flegontov, P.; Altınışık, N.E.; Changmai, P.; Rohland, N.; Mallick, S.; Adamski, N.; Bolnick, D.A.; Broomandkhoshbacht, N.; Candilio, F.; Cullen, B.J.; et al. Palaeo-Eskimo genetic ancestry and the peopling of Chukotka and North America. *Nature* **2019**, *570*, 236–240. [\[CrossRef\]](#)

86. Gnecci-Ruscione, G.A.; Sarno, S.; De Fanti, S.; Gianvincenzo, L.; Giuliani, C.; Boattini, A.; Bortolini, E.; di Corcia, T.; Mellado, C.S.; Francia, T.J.D.; et al. Dissecting the Pre-Columbian Genomic Ancestry of Native Americans along the Andes–Amazonia Divide. *Mol. Biol. Evol.* **2019**, *36*, 1254–1269. [\[CrossRef\]](#)
87. Reynolds, A.W.; Mata-Míguez, J.; Miró-Herrans, A.; Briggs-Cloud, M.; Sylestine, A.; Barajas-Olmos, F.; Garcia-Ortiz, H.; Rzhetskaya, M.; Orozco, L.; Raff, J.A.; et al. Comparing signals of natural selection between three Indigenous North American populations. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 9312–9317. [\[CrossRef\]](#)
88. Barbieri, C.; Barquera, R.; Arias, L.; Sandoval, J.R.; Acosta, O.; Zurita, C.; Aguilar-Campos, A.; Tito-Álvarez, A.M.; Serrano-Osuna, R.; Gray, R.D.; et al. The Current Genomic Landscape of Western South America: Andes, Amazonia, and Pacific Coast. *Mol. Biol. Evol.* **2019**, *36*, 2698–2713. [\[CrossRef\]](#)
89. Sánchez-Pozos, K.; Ortiz-López, M.G.; Peña-Espinoza, B.I.; de los Ángeles Granados-Silvestre, M.; Jiménez-Jacinto, V.; Verleyen, J.; Tekola-Ayele, F.; Sanchez-Flores, A.; Menjivar, M. Whole-exome sequencing in maya indigenous families: Variant in PPP1R3A is associated with type 2 diabetes. *Mol. Genet. Genom.* **2018**, *293*, 1205–1216. [\[CrossRef\]](#) [\[PubMed\]](#)
90. Moreno-Mayar, J.V.; Potter, B.A.; Vinner, L.; Steinrücken, M.; Rasmussen, S.; Terhorst, J.; Kamm, J.; Albrechtsen, A.; Malaspinas, A.-S.; Sikora, M.; et al. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature* **2018**, *553*, 203–207. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Harris, D.N.; Song, W.; Shetty, A.C.; Levano, K.S.; Cáceres, O.; Padilla, C.; Borda, V.; Tarazona, D.; Trujillo, O.; Sanchez, C.; et al. Evolutionary genomic dynamics of Peruvians before, during, and after the Inca Empire. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E6526–E6535. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Lindo, J.; Haas, R.; Hofman, C.; Apata, M.; Moraga, M.; Verdugo, R.A.; Watson, J.T.; Llave, C.V.; Witonsky, D.; Beall, C.; et al. The genetic prehistory of the Andean highlands 7000 years BP though European contact. *Sci. Adv.* **2018**, *4*, eaau4921. [\[CrossRef\]](#)
93. Schroeder, H.; Sikora, M.; Gopalakrishnan, S.; Cassidy, L.M.; Maisano Delser, P.; Sandoval Velasco, M.; Schraiber, J.G.; Rasmussen, S.; Homburger, J.R.; Ávila-Arcos, M.C.; et al. Origins and genetic legacies of the Caribbean Taino. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2341–2346. [\[CrossRef\]](#)
94. de la Fuente, C.; Ávila-Arcos, M.C.; Galimany, J.; Carpenter, M.L.; Homburger, J.R.; Blanco, A.; Contreras, P.; Dávalos, D.C.; Reyes, O.; San Roman, M.; et al. Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E4006–E4012. [\[CrossRef\]](#)
95. Moreno-Mayar, J.V.; Vinner, L.; de Barros Damgaard, P.; de la Fuente, C.; Chan, J.; Spence, J.P.; Allentoft, M.E.; Vimala, T.; Racimo, F.; Pinotti, T.; et al. Early human dispersals within the Americas. *Science* **2018**, *362*, eaav2621. [\[CrossRef\]](#)
96. Posth, C.; Nakatsuka, N.; Lazaridis, I.; Skoglund, P.; Mallick, S.; Lamnidis, T.C.; Rohland, N.; Nägele, K.; Adamski, N.; Bertolini, E.; et al. Reconstructing the Deep Population History of Central and South America. *Cell* **2018**, *175*, 1185–1197.e22. [\[CrossRef\]](#)
97. Peng, Q.; Schork, N.J.; Wilhelmsen, K.C.; Ehlers, C.L. Whole genome sequence association and ancestry-informed polygenic profile of EEG alpha in a Native American population. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **2017**, *174*, 435–450. [\[CrossRef\]](#)
98. Romero-Hidalgo, S.; Ochoa-Leyva, A.; Garcíarrubio, A.; Acuña-Alonzo, V.; Antúnez-Argüelles, E.; Balcazar-Quintero, M.; Barquera-Lozano, R.; Carnevale, A.; Cornejo-Granados, F.; Fernández-López, J.C.; et al. Demographic history and biologically relevant genetic variation of Native Mexicans inferred from whole-genome sequencing. *Nat. Commun.* **2017**, *8*, 1005. [\[CrossRef\]](#)
99. Lindo, J.; Achilli, A.; Perego, U.A.; Archer, D.; Valdiosera, C.; Petzelt, B.; Mitchell, J.; Worl, R.; Dixon, E.J.; Fifield, T.E.; et al. Ancient individuals from the North American Northwest Coast reveal 10,000 years of regional genetic continuity. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 4093–4098. [\[CrossRef\]](#)
100. Fehren-Schmitz, L.; Jarman, C.L.; Harkins, K.M.; Kayser, M.; Popp, B.N.; Skoglund, P. Genetic Ancestry of Rapanui before and after European Contact. *Curr. Biol.* **2017**, *27*, 3209–3215.e6. [\[CrossRef\]](#)
101. Lindo, J.; Huerta-Sánchez, E.; Nakagome, S.; Rasmussen, M.; Petzelt, B.; Mitchell, J.; Cybulski, J.S.; Willerslev, E.; DeGiorgio, M.; Malhi, R.S. A time transect of exomes from a Native American population before and after European contact. *Nat. Commun.* **2016**, *7*, 13175. [\[CrossRef\]](#)
102. Manousaki, D.; Kent, J.W., Jr.; Haack, K.; Zhou, S.; Xie, P.; Greenwood, C.M.; Brassard, P.; Newman, D.E.; Cole, S.; Umans, J.G.; et al. Toward Precision Medicine: TBC1D4 Disruption Is Common Among the Inuit and Leads to Underdiagnosis of Type 2 Diabetes. *Diabetes Care* **2016**, *39*, 1889–1895. [\[CrossRef\]](#)
103. Dewey, F.E.; Murray, M.F.; Overton, J.D.; Habegger, L.; Leader, J.B.; Fetterolf, S.N.; O’Dushlaine, C.; van Hout, C.V.; Staples, J.; Gonzaga-Jauregui, C.; et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science* **2016**, *354*, aaf6814. [\[CrossRef\]](#)
104. Mallick, S.; Li, H.; Lipson, M.; Mathieson, I.; Gymrek, M.; Racimo, F.; Zhao, M.; Chennagiri, N.; Nordenfelt, S.; Tandon, A.; et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **2016**, *538*, 201–206. [\[CrossRef\]](#)
105. Zhou, S.; Xiong, L.; Xie, P.; Ambalavanan, A.; Bourassa, C.V.; Dionne-Laporte, A.; Spiegelman, D.; Gauthier, M.T.; Henrion, E.; Diallo, O.; et al. Increased Missense Mutation Burden of Fatty Acid Metabolism Related Genes in Nunavik Inuit Population. *PLoS ONE* **2015**, *10*, e0128255. [\[CrossRef\]](#)
106. Rasmussen, M.; Sikora, M.; Albrechtsen, A.; Korneliussen, T.S.; Moreno-Mayar, J.V.; Poznik, G.D.; Zollikofer, C.P.E.; de León, M.S.P.; Allentoft, M.; Moltke, I.; et al. The ancestry and affiliations of Kennewick Man. *Nature* **2015**, *523*, 455–458. [\[CrossRef\]](#)
107. Skoglund, P.; Mallick, S.; Bortolini, M.C.; Chennagiri, N.; Hünemeier, T.; Petzl-Erler, M.L.; Salzano, F.M.; Patterson, N.; Reich, D. Genetic evidence for two founding populations of the Americas. *Nature* **2015**, *525*, 104–108. [\[CrossRef\]](#)

108. Malaspinas, A.S.; Lao, O.; Schroeder, H.; Rasmussen, M.; Raghavan, M.; Moltke, I.; Campos, P.; Sagredo, F.S.; Rasmussen, S.; Gonçalves, V.F.; et al. Two ancient human genomes reveal Polynesian ancestry among the indigenous Botocudos of Brazil. *Curr. Biol.* **2014**, *24*, R1035–R1037. [\[CrossRef\]](#)
109. Raghavan, M.; DeGiorgio, M.; Albrechtsen, A.; Moltke, I.; Skoglund, P.; Korneliussen, T.S.; Grønnow, B.; Appelt, M.; Gulløv, H.C.; Friesen, T.M.; et al. The genetic prehistory of the New World Arctic. *Science* **2014**, *345*, 1255832. [\[CrossRef\]](#)
110. Rasmussen, M.; Anzick, S.L.; Waters, M.R.; Skoglund, P.; DeGiorgio, M.; Stafford, T.W.; Rasmussen, S.; Moltke, I.; Albrechtsen, A.; Doyle, S.M.; et al. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **2014**, *506*, 225–229. [\[CrossRef\]](#)
111. Verdu, P.; Pemberton, T.J.; Laurent, R.; Kemp, B.M.; Gonzalez-Oliver, A.; Gorodezky, C.; Hughes, C.E.; Shattuck, M.R.; Petzelt, B.; Mitchell, J.; et al. Patterns of Admixture and Population Structure in Native Populations of Northwest North America. *PLoS Genet.* **2014**, *10*, e1004530. [\[CrossRef\]](#)
112. Moreno-Estrada, A.; Gignoux, C.R.; Fernández-López, J.C.; Zakharia, F.; Sikora, M.; Contreras, A.V.; Acuña-Alonzo, V.; Sandoval, K.; Eng, C.; Romero-Hidalgo, S.; et al. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* **2014**, *344*, 1280–1285. [\[CrossRef\]](#)
113. Moreno-Mayar, J.V.; Rasmussen, S.; Seguin-Orlando, A.; Rasmussen, M.; Liang, M.; Flåm, S.T.; Lie, B.A.; Gilfillan, G.D.; Nielsen, R.; Thorsby, E.; et al. Genome-wide Ancestry Patterns in Rapanui Suggest Pre-European Admixture with Native Americans. *Curr. Biol.* **2014**, *24*, 2518–2525. [\[CrossRef\]](#)
114. Lazaridis, I.; Patterson, N.; Mittnik, A.; Renaud, G.; Mallick, S.; Kirsanow, K.; Sudmant, P.H.; Schraiber, J.G.; Castellano, S.; Lipson, M.; et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **2014**, *513*, 409–413. [\[CrossRef\]](#)
115. Ribeiro-dos-Santos, A.M.; de Souza, J.E.S.; Almeida, R.; Alencar, D.O.; Barbosa, M.S.; Gusmão, L.; Silva, W.A., Jr.; de Souza, S.J.; Silva, W.A.; Darnet, S.; et al. High-Throughput Sequencing of a South American Amerindian. *PLoS ONE* **2013**, *8*, e83340. [\[CrossRef\]](#)
116. Moreno-Estrada, A.; Gravel, S.; Zakharia, F.; McCauley, J.L.; Byrnes, J.K.; Gignoux, C.R.; Ortiz-Tello, P.A.; Martínez, R.J.; Hedges, D.J.; Morris, R.W.; et al. Reconstructing the Population Genetic History of the Caribbean. *PLoS Genet.* **2013**, *9*, e1003925. [\[CrossRef\]](#)
117. Reich, D.; Patterson, N.; Campbell, D.; Tandon, A.; Mazieres, S.; Ray, N.; Parra-Marín, M.V.; Rojas, W.; Duque, C.; Mesa, N.; et al. Reconstructing Native American population history. *Nature* **2012**, *488*, 370–374. [\[CrossRef\]](#)
118. Patterson, N.; Moorjani, P.; Luo, Y.; Mallick, S.; Rohland, N.; Zhan, Y.; Genschoreck, T.; Webster, T.; Reich, D. Ancient Admixture in Human History. *Genetics* **2012**, *192*, 1065–1093. [\[CrossRef\]](#)
119. Reich, D.; Green, R.E.; Kircher, M.; Krause, J.; Patterson, N.; Durand, E.Y.; Viola, B.; Briggs, A.W.; Stenzel, U.; Johnson, P.L.F.; et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **2010**, *468*, 1053–1060. [\[CrossRef\]](#) [\[PubMed\]](#)
120. Rasmussen, M.; Li, Y.; Lindgreen, S.; Pedersen, J.S.; Albrechtsen, A.; Moltke, I.; Metspalu, M.; Metspalu, E.; Kivisild, T.; Gupta, R.; et al. Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **2010**, *463*, 757–762. [\[CrossRef\]](#) [\[PubMed\]](#)
121. Silva-Zolezzi, I.; Hidalgo-Miranda, A.; Estrada-Gil, J.; Fernandez-Lopez, J.C.; Uribe-Figueroa, L.; Contreras, A.; Balam-Ortiz, E.; del Bosque-Plata, L.; Velazquez-Fernandez, D.; Lara, C.; et al. Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8611–8616. [\[CrossRef\]](#) [\[PubMed\]](#)
122. Byrd, J.B.; Greene, A.C.; Prasad, D.V.; Jiang, X.; Greene, C.S. Responsible, practical genomic data sharing that accelerates research. *Nat. Rev. Genet.* **2020**, *21*, 615–629. [\[CrossRef\]](#)
123. Wilson, S.L.; Way, G.P.; Bittremieux, W.; Armache, J.P.; Haendel, M.A.; Hoffman, M.M. Sharing biological data: Why, when, and how. *FEBS Lett.* **2021**, *595*, 847–863. [\[CrossRef\]](#)
124. Crampton, P.; Parkin, C. Warrior genes and risk-taking science. *N. Z. Med. J.* **2007**, *120*, U2439.
125. Hudson, M.; Garrison, N.A.; Sterling, R.; Caron, N.R.; Fox, K.; Yracheta, J.; Anderson, J.; Wilcox, P.; Arbour, L.; Brown, A.; et al. Rights, interests and expectations: Indigenous perspectives on unrestricted access to genomic data. *Nat. Rev. Genet.* **2020**, *21*, 377–384. [\[CrossRef\]](#)
126. Alpaslan-Roodenberg, S.; Anthony, D.; Babiker, H.; Bánffy, E.; Booth, T.; Capone, P.; Deshpande-Mukherjee, A.; Eisenmann, S.; Fehren-Schmitz, L.; Frachetti, M.; et al. Ethics of DNA research on human remains: Five globally applicable guidelines. *Nature* **2021**, *599*, 41–46. [\[CrossRef\]](#)
127. Séguin, B.; Hardy, B.J.; Singer, P.A.; Daar, A.S. Genomics, public health and developing countries: The case of the Mexican National Institute of Genomic Medicine (INMEGEN). *Nat. Rev. Genet.* **2008**, *9*, S5–S9. [\[CrossRef\]](#)
128. Hetu, M.; Koutouki, K.; Joly, Y. Genomics for All: International Open Science Genomics Projects and Capacity Building in the Developing World. *Front. Genet.* **2019**, *10*, 95. [\[CrossRef\]](#)
129. Séguin, B.; Hardy, B.J.; Singer, P.A.; Daar, A.S. Genomic medicine and developing countries: Creating a room of their own. *Nat. Rev. Genet.* **2008**, *9*, 487–493. [\[CrossRef\]](#)
130. Helmy, M.; Awad, M.; Mosa, K.A. Limited resources of genome sequencing in developing countries: Challenges and solutions. *Appl. Transl. Genom.* **2016**, *9*, 15–19. [\[CrossRef\]](#)
131. Meagher, K.M.; Lee, L.M. Integrating Public Health and Deliberative Public Bioethics: Lessons from the Human Genome Project Ethical, Legal, and Social Implications Program. *Public Health Rep.* **2016**, *131*, 44. [\[CrossRef\]](#)
132. de Vries, J.; Bull, S.J.; Doumbo, O.; Ibrahim, M.; Mercereau-Puijalon, O.; Kwiatkowski, D.; Parker, M. Ethical issues in human genomics research in developing countries. *BMC Med. Ethics* **2011**, *12*, 5. [\[CrossRef\]](#)

133. Mello, M.M.; Wolf, L.E. The Havasupai Indian Tribe Case—Lessons for Research Involving Stored Biologic Samples. *N. Engl. J. Med.* **2010**, *363*, 204–207. [CrossRef]
134. Dalton, R. Tribe blasts “exploitation” of blood samples. *Nature* **2002**, *420*, 111. [CrossRef]
135. Merriman, T.; Cameron, V. Risk-taking: Behind the warrior gene story. *N. Z. Med. J.* **2007**, *120*, U2440.
136. The International Work Group for Indigenous Affairs (IWGIA). *The Indigenous World 2021*, 35th ed.; IWGIA: Copenhagen, Denmark, 2021; 824p, Available online: https://iwgia.org/doclink/iwgia-book-the-indigenous-world-2021-eng/eyJ0eXAiOiJKV1QiLCJhbGciOiJIUzI1NiJ9.eyJzdWIiOiJpd2dpYS1ib29rLXRoZS1pbmRpZ2Vub3VzLXdvcmxkLTlwMjEtZW5nliwiaWF0IjoxNjI4ODM5NjM2LCJleHAiOiJlE2Mjg5MjYwMzZ9.z1CuM7PcT5CPkV0evx8ve88y6v0vmwDu_51JQ_lwAkM (accessed on 22 March 2022).
137. Schauburger, P.; Walker, A.; Braglia, L.; Sturm, J.; Garbuszus, J.M.; Barbone, J.M. Openxlsx: Read, Write and Edit xlsx Files. 2021. Available online: <https://CRAN.R-project.org/package=openxlsx> (accessed on 11 March 2022).
138. Wickham, H.; François, R.; Henry, L.; Müller, K.; RStudio. Dplyr: A Grammar of Data Manipulation. 2022. Available online: <https://CRAN.R-project.org/package=dplyr> (accessed on 11 March 2022).
139. Wickham, H.; Chang, W.; Henry, L.; Pedersen, T.L.; Takahashi, K.; Wilke, C.; Woo, K.; Yutani, H.; Dunnington, D.; RStudio. Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. 2021. Available online: <https://CRAN.R-project.org/package=ggplot2> (accessed on 11 March 2022).
140. Slowikowski, K.; Schep, A.; Hughes, S.; Dang, T.K.; Lukauskas, S.; Irisson, J.-O.; Kamvar, Z.N.; Ryan, T.; Christophe, D.; Hiroaki, Y.; et al. Ggrepel: Automatically Position Non-Overlapping Text Labels with “ggplot2”. 2021. Available online: <https://CRAN.Rproject.org/package=ggrepel> (accessed on 11 March 2022).
141. Wilkins, D.; Rudis, B. Treemapify: Draw Treemaps in “ggplot2”. 2021. Available online: <https://CRAN.R-project.org/package=treemapify> (accessed on 11 March 2022).
142. Wilke, C.O. Cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2”. 2020. Available online: <https://CRAN.Rproject.org/package=cowplot> (accessed on 11 March 2022).
143. Xiao, N.; Li, M. Ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for “ggplot2”. 2018. Available online: <https://CRAN.Rproject.org/package=ggsci> (accessed on 11 March 2022).
144. Wickham, H.; Girlich, M.; RStudio. Tidy: Tidy Messy Data. 2022. Available online: <https://CRAN.R-project.org/package=tidy> (accessed on 11 March 2022).
145. Wickham, H.; Seidel, D.; RStudio. Scales: Scale Functions for Visualization. 2020. Available online: <https://CRAN.R-project.org/package=scales> (accessed on 11 March 2022).
146. Xie, Y. Knitr: A General-Purpose Package for Dynamic Report Generation in R. 2021. Available online: <https://yihui.org/knitr/> (accessed on 11 March 2022).
147. Xie, Y. *Dynamic Documents with R and Knitr*, 2nd ed.; Chapman and Hall: London, UK; CRC: Boca Raton, FL, USA, 2015; Available online: <https://www.routledge.com/Implementing-Reproducible-Research/Stodden-Leisch-Peng/p/book/9780367576172> (accessed on 11 March 2022).
148. Stodden, V.; Leisch, F.; Peng, R.D. *Implementing Reproducible Research*; CRC Press: Boca Raton, FL, USA, 2014; 440p.
149. South, A. Rnaturalearth: World Map Data from Natural Earth. 2017. Available online: <https://CRAN.R-project.org/package=rnaturalearth> (accessed on 11 March 2022).
150. Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *R J.* **2018**, *10*, 439–446. [CrossRef]
151. Tennekes, M. tmap: Thematic Maps in R. *J. Stat. Softw.* **2018**, *84*, 1–39. [CrossRef]
152. Bivand, R.; Rundel, C.; Pebesma, E.; Stuetz, R.; Hufthammer, K.O.; Giraudoux, P.; Davis, M.; Santilli, S. Rgeos: Interface to Geometry Engine—Open Source (‘GEOS’). 2021. Available online: <https://CRAN.R-project.org/package=rgeos> (accessed on 11 March 2022).